МОДИФИЦИРОВАННЫЙ ГИБРИДНЫЙ ГЕНЕТИЧЕСКИЙ АЛГОРИТМ РЕШЕНИЯ ЗАДАЧ ДИСКРЕТНОЙ ОПТИМИЗАЦИИ

А.О. Глухов[1], В.В. Трофимов[2], Д.О. Глухов[1], Л.А. Трофимова[2]

[1]*Полоцкий державный университет*
[2]*Санкт-Петербургский государственный экономический университет*

Abstract. *The goal objective* is to improve the efficiency of solving discrete optimization problems. The proposed method refers to the "fast" methods and was named the "Local genetic method". The peculiarity of this method is that the chromosomes do not encode the whole solution, but only a small part of the plan. Therefore, the method allows us introducing unary and binary operations that take into account the specific nature of the problem. The important feature of the method is the non-deterministic nature of the computation, which is due to the internal parallelism of computations and is expressed in the asynchronous action of various local strategies. In terms of speed, the proposed method in a number of experiments outperformed the traditional algorithm by more than 10 times and always found the best solution. The nature of the approximation to the optimum for these algorithms remained unchanged when solving any test cases.

Keywords: genetic algorithms; discrete optimization

Задача планирования (расписания, маршрута коммивояжера, многопрофильного производства и др.), в целом, характеризуется чрезвычайно высокой временной сложностью и относится к классу задач дискретной оптимизации. Она эффективно решается только с применением приближенных методов, таких как генетический поиск или эвристический метод. Предлагаемый метод также относится к «быстрым» методам и получил название «Локальный генетический метод» (ЛГМ) [1]. Метод можно классифицировать как модифицированный гибридный генетический алгоритм. Особенностью метода является то, что хромосомы кодируют не целое решение, а лишь небольшой участок плана. В связи с чем, метод не использует традиционные операции над хромосомами (кроссинговер, инверсию и т.п.), а позволяет вводить унарные и бинарные операции, учитывающие специфику задачи. Такие операции имеют чисто эвристический характер, и их удачный выбор позволяет достичь максимальной эффективности процесса поиска решения в задаче планирования.

Для иллюстрации работы ЛГМ была выбрана задача коммивояжера (traveling salesman problem - TSP), так как данная задача достаточно широко освещена в литературе и существуют тестовые наборы примеров, пользуясь которыми мы можем определить показатели сравнительной эффективности нового метода с существующими методами, такими как: эвристические методы, оптимизация муравьиными колониями, генетические методы, «табу»-поиск, нейронные сети, аппроксимация и т.д.

В основе генетического алгоритма лежит компьютерная модель эволюции [2,3,5,6]. Каждая особь характеризуется своей хромосомой, представленной в виде цепочки символов:

$$X=\{s_1,s_2,\ldots,s_g|s_i\in SYMBOL\}\in Chromosome.$$

Символы выбираются из множества допустимых символов (генов) – SYMBOL, кодирующих узлы (города). При этом длина хромосом может изменяться в пределах $[1,\ldots,|SYMBOL|]$. Мы будем пользоваться традиционной для генетического алгоритма терминологией, несмотря на различия подходов. Задачу оптимизации обычно формулируют как максимизацию функции приспособленности особи $f(X)$, т.е.

$$\max_{X\in Chromosome} f(X) \quad \Big|\quad f(X):\ Chromosome\to R^+,$$

где R^+ – множество неотрицательных действительных чисел.

Для эвристического алгоритма характерно использование другой функции для сравнения качества решений – функции предпочтения $Q(X)$. Задача же оптимизации, в этом случае, формулируется как $Q(X)\to max$. Функция предпочтения обычно строится как эвристическая функция штрафов за несовершенство решения. На самом деле, нет никаких препятствий, чтобы перейти к использованию эвристической функции предпочтения вместо функции приспособленности в генетическом алгоритме. При этом мы получаем возможность оперировать также и «неполными» хромосомами.

Для осуществления поиска нам необходимо иметь определенный набор операций над хромосомами $ACTION=\{a_1,a_2,...,a_M\}$ и определенным набором условий для выполнения этих операций $CONDITION=\{c_1,c_2,...,c_M\}$.

Эвристической стратегией σ мы будем называть пару, состоящую из операции и условия ее активизации:

$\sigma_i=(a_i,c_i)\in STRATEGY$.

c_i: $Chromosomen^n\to B|\ n=1,2,...$ – предикат, где $B=\{true,false\}$ – пространство логических переменных.

a_i: $Chromosomen^n\to Chromosome\ |\ n=1,2,...$ – унарные, бинарные и n-арные операции.

Вычисления с использованием вышеуказанных эвристических стратегий требуют задаться некоторой популяцией $P\subseteq Chromosome$. До начала вычислений мы должны задаться начальной популяцией P_0, которая, например, может строится случайным образом.

Формализуем локальный генетический алгоритм решения задачи TSP:

```
LGA: 1. создать начальную популяцию P0;
     2. вычислить Q(X), X∈Pk, где Q – функция предпочтения;
     3. сформировать популяцию следующего поколения Pk+1=Sep(Pk∪Pk/), где Sep – оператор отбора, Pk/ - множество особей, полученное вследствие действия набора стратегий STRATEGY;
     4. если не выполняется условие окончания эволюционного поиска – перейти к пункту 2;
```

Очевидным преимуществом LGA является снижение вычислительной сложности алгоритма за счет уменьшения длины подвергаемых анализу и изменению цепочек, а также за счет того факта, что для эвристических операций, как правило, удается вести расчет изменения значения функции предпочтения, а не вычислять ее новое значение «с нуля». Другим преимуществом LGA является повышенная гибкость по сравнению с классическим вариантом генетического алгоритма, которая выражается в совместном действии различных стратегий, а также в возможности легко видоизменять участвующие в нем эвристические стратегии (вводить новые, включать или отключать их).

Важной особенностью LGA является недетерминистический характер вычислений, который является следствием внутреннего параллелизма вычислений. Последний выражается в асинхронности действия различных локальных стратегий. Отсутствие искусственной случайности позволяет двигаться в пространстве решений задачи целенаправленно, совершая шаги, только в сторону «*возможного улучшения*» значения функции предпочтения (так мы понимаем эвристическую стратегию). Так как возможности значительного «*улучшения*» уменьшаются при приближении к *оптимуму*, то характер движения в LGA становится более предсказуемым, по сравнению с традиционным генетическим алгоритмом.

Основным недостатком алгоритма является его эвристический характер. Сочетание различных стратегий позволяет преодолеть большее число локальных минимумов, по сравнению с обычным эвристическим алгоритмом, но *не гарантирует* достижение оптимума. Второй очевидный недостаток (он является следствием первого) – мы можем располагать лишь статистическими оценками эффективности данного алгоритма, что требует проведения огромного количества экспериментов на различных задачах для получения достоверной информации.

Выбор стратегий. В задаче TSP хромосома кодирует маршрут движения коммивояжера, который мы можем оценить с помощью следующей функции предпочтения:

$Q(X) = Length + A\cdot(|X|-N) + B\cdot DoNotClose$,

где Length – длина маршрута при обязательном возвращении коммивояжера в пункт отправления, A – штраф за пропуск города, B – штраф за незавершенность маршрута, DoNotClose – равен 1, если S_i – частичное решение, иначе 0. Использованы следующие значения для штрафных переменных (A=1000, B=3000) для модели географического региона размером ~ 500x400.

Для формализации описания эвристических стратегий введем оператор порождения новой особи (хромосомы) – $X_{PARENT} \Rightarrow X_{NEW}$. Множество эвристических стратегий STRATEGY будет содержать следующие стратегии $\{\sigma_{grow}, \sigma_{transition}, \sigma_{exchange}, \sigma_{rotation}\}$. Далее для всех стратегий справедливо: $X, X' \in Chromosome$, $COND = s_x \notin X \wedge s_x \in SYMBOL$.

Ниже формально определены стратегии:

$\sigma_{grow} = ((X \Rightarrow X' = X \cup \langle s_x \rangle), COND)$.

$\sigma_{transition} = ((X = X_B \cup \langle s_i \rangle \cup X_M \cup X_E \Rightarrow X' = X_B \cup X_M \cup \langle s_i \rangle \cup X_E), true)$.

$\sigma_{exchange} = ((X = X_B \cup \langle s_i \rangle \cup X_M \cup \langle s_m \rangle \cup X_E \Rightarrow X' = X_B \cup \langle s_m \rangle \cup X_M \cup \langle s_i \rangle \cup X_E), true)$.

$\sigma_{rotation} = ((X = X_B \cup \langle s_i \rangle \cup X_M \cup \langle s_m \rangle \cup X_E \Rightarrow X' = X_B \cup \langle s_i \rangle \cup inverce(X_M) \cup \langle s_m \rangle \cup X_E), true)$.

Функция inverce(X) изменяет порядок следования символов в наборе X на обратный. Все стратегии позволяют произвести расчет изменения значения функции предпочтения для хромосом, что значительно снижает вычислительные затраты.

Результаты. Сравнительные показатели эффективности традиционного генетического алгоритма - GA и LGA приводятся ниже в таблице 1. Для постановки экспериментов подбирались параметры алгоритмов, при которых удавалось достичь наилучших результатов на одних и тех же тестовых примерах. Так, например, для GA – размер популяции - 50, элитное скрещивание - 50, мутация - 3, инверсия - 8; для LGA – размер популяции - 50, переход - 6, кручение - 6, обмен - 6. Условием прекращения поиска являлось отсутствие заметных улучшения решения в течении определенного времени (около 15% времени вычислений).

Таблица 1.

Size	GA		LGA	
	t, sec	Q(X)	t, sec	Q(X)
50	12	2288	3	2076
80	28	4408	4	2930
140	35	4441	12	4142
200	135	5103	17	4616
270	261	4755	36	4431
350	710	6021	51	5498

Источник: Программа LGA.exe.

Проводились отдельные эксперименты и на тестовых примерах, взятых из TSPLIB. Результаты расчетов приведены в табл.2.

Таблица 2.

Task	Size	current optimum (TSPLIB)	LGA	
		Q(X)	t, sec	Q(X)
eil51	51	429.983	3:04	428.982
rd100	100	7909.1	10:32	7908.04
eil76	76	545.388	4:14	545.388
pr76	76	108159	17:12	108159
cp150	150	6527.85	30:40	6566.58

В первых двух задачах нам удалось улучшить решение, которое считалось оптимальным до настоящего момента.

Заключение

Полученные результаты свидетельствуют в пользу предложенных алгоритмов. Скорость и точность заметно превышает соответствующие характеристики традиционного алгоритма. Так по скорости LGA в ряде экспериментов опередил традиционный алгоритм более чем в 10 раз и всегда находил лучшее решение. «Разбежка» в качестве решений в среднем составила около 9%-12%. Характер приближения к оптимуму для этих алгоритмов оставался неизменным при решении любых тестовых примеров.

Интересен также и тот факт, что мы полностью исключили из генетического алгоритма бинарные операции (например, *скрещивание*) и при этом его эффективность возросла за счет введения достаточного количества унарных (*мутационных*) операций специального вида. Последние, как правило, позволяют оценить величину изменения цены решения ΔQ, что устраняет необходимость полного расчета цены новых решений. Так для $X^{/}=a_i(X)$ имеем $Q(X^{/})=Q(X)+\Delta Q$. Таким образом, вычислительные затраты новых алгоритмов становятся заметно ниже. Это позволяет нам предполагать, что эффективность локальных генетических алгоритмов будет значительно превышать эффективность традиционного алгоритма именно при решении задач большой и очень большой размерности. Таковыми являются, например, задачи планирования и составления расписаний, транспортная задача с большим количеством «*городов*» и т.д.

В качестве направления дальнейших исследований мы видим совершенствование стратегий поиска в LGA и, возможно, внесение адаптивного поведения в алгоритм.

Список литературы

1. Глухов А.О., Трофимов В.В., Трофимова Л.А. Аппроксимация сложных, многопараметрических, существенно нелинейных, динамических зависимостей на основе генетических алгоритмов. XIX Международная конференция по мягким вычислениям и измерениям (SCM-2016). Сборник докладов в 2-х томах. Санкт-Петербург. 25-27 мая 2016 г. Т.1. – С.477-480.

2. Глухов А.О., Трофимов В.В. Использование рекурсивных эвристик при решении задач дискретной оптимизации большой размерности // Современные проблемы менеджмента: межвуз. сб. Выпуск 6. – СПб.: Изд-во СПбГУЭФ, 2003. – С. 99-105.

3. Глухов А.О., Трофимов В.В., Глухов Д.О. Локальный генетический алгоритм планирования процесса многопрофильного производства / Экономическая кибернетика: системный анализ в экономике и управлении: Сб. научных трудов. Выпуск №5 – СПб.: СПбГУЭФ, 2002 – С. 50-56.

4. Trofimov V.V., Hluhov A.O. The Optimal Schedule for the Technological Process of Semiconductor Production. Operations Research 2002: International Conference on Operations Research Sept.2-sept.5, 2002, Klagenfurt, Austria, Klagenfurt University, 2002. – P.148.

5. Трофимов В.В., Глухов А.О. Метод решения задачи идентификации и моделирования самоорганизующихся систем на основе генетического подхода. Международная конференция по мягким вычислениям и измерениям, Санкт-Петербург, 25-27 июня 2001, Сборник докладов. СПб.: ЛЭТИ, 2001, том 1. – с.291-293.

6. Trofimov V.V., Hluhov A.O. Construction of the Optimal Schedule for the Technological process of semiconducting production. 46. International Scientific Colloquium. “Multimedia – The Challenge for Science, Technology and Business”. 24-27.09.2001 Technique University of Ilmenau, Germany 2001.